

引言

2019年2月2日 14:18



◆ 数据结构

1. 数据是对客观事物的符号表示，在计算机科学中指所有能输入进计算机并被计算机程序处理的符号的总称
 - 1) 数据元素data element是数据的基本单位，在计算机程序中通常被作为一个整体进行考虑和处理，一般由若干数据项data item组成
 - 2) 数据项是数据不可分割的最小单位
 - 3) 数据对象data object是性质相同的数据元素的集合，是数据的子集
2. 数据结构是相互之间存在一种或多种特定关系的数据元素的集合
 - 1) 数据结构是一个二元组，记为 $data_structure=(D,S)$ ，其中D为数据元素的集合，S是D上关系（逻辑结构）的集合
 - 2) 数据结构是一门研究非数值计算的程序设计问题中计算机的操作对象以及它们之间的关系和操作等的学科
3. 数据元素之间的逻辑结构关系
 - 1) 集合关系：各数据元素同属于一个集合
 - 2) 线性结构：数据元素间存在顺序关系
 - 3) 树状结构：结构中的元素间的关系是一对多的关系
 - 4) 图(网)状结构：结构中的元素间的关系是多对多的关系
4. 存储方式和逻辑关系没有必然相同的关系
 - 1) 逻辑结构：数据元素之间存在的关系（逻辑关系）称为数据的逻辑结构
 - 2) 物理结构：数据结构在计算机中的表示（映象）称为数据的物理结构
 - 3) 一种逻辑结构可映象成不同的存储结构
5. 抽象数据类型是指一个数学模型以及定义在此数学模型上的一组操作
 - 1) 抽象数据类型=数据结构+定义在此数据结构上的一组操作
 - 2) 可用 (D, S, P) 三元组表示。其中，D是数据对象，S是D上的关系集，P是对D的基本操作集
 - 3) ADT 抽象数据类型名 {
 数据对象：〈数据对象的定义〉
 数据关系：〈数据关系的定义〉
 基本操作：〈基本操作的定义〉
} ADT 抽象数据类型名
 - i. 数据对象、数据关系用伪码描述
 - ii. 基本操作定义格式为
 基本操作名（参数表）
 初始条件：〈初始条件描述〉
 操作结果：〈操作结果描述〉



◆ 算法

6. 算法就是对特定的数据结构进行一系列操作，其目的是解决一个具体问题

1) 算法 + 数据结构 = 程序

7. 运算序列四大特征

1) 有穷性，即序列的项数有限，且每一运算项都可在有限的时间内完成

2) 确定性，即序列的每一项运算都有明确的定义，无二义性

3) 输入输出：可以没有输入运算项，但一定要有输出运算项

4) 可行性，即对于任意给定的合法的输入都能得到相应的正确的输出

8. 算法设计的要求

1) 正确性correctness：无语法错、对普通数据、刁难数据、一切合法输入数据都能得到符合要求的结果

2) 可读性readability：有助于人的理解

3) 健壮性robustness：对非法输入能做出反应

4) 效率与低存储量需求：执行时间短、所需最大存储空间小

9. 算法复杂度Complexity(C)

1) 时间复杂度T(Time)：算法运行所需要的时间资源的量

2) 空间复杂度S(Space)：算法运行所需要的空间资源的量（包括内存和外存）

3) 依赖于：问题的规模；算法的输入；算法本身

i. 简写为 $C = f(N, I, A)$ Dimensions(N)、Input(I)和A(Algorism)来表示

4) 时间复杂度计算主要考虑算法中最高阶项带来的开销。渐近时间复杂度表示：

$T(n)=O(f(n))$

i. “O”定义：“若 $T(n)$ 和 $f(n)$ 是定义在正整数集合上的两个函数，则 $T(n)=O(f(n))$ 表示存在正的常数 C 和 n_0 ,使得当 $n \geq n_0$ 时都满足 $0 \leq T(n) \leq C * f(n)$ ”。即这两个函数当整型自变量 n 趋向于无穷大时，两者的比值是一个不等于0的常数

ii. 当要计算某个算法的时间复杂度 $T(N)$ 时，找一个更简单的，阶数相同的简单算法 f ，可以等同计算，因此可以记为： $T(N)=O(f(N))$ 这里的 f 指的是那个替代函数，它具有和原算法一样最高阶复杂度

iii. 对于算法的复杂度，我们只考虑最坏、最好和平均三种情况，而通常又着重于最坏情况。本节只对最坏情况进行讨论

5) 分析时间复杂度的七条规则基本覆盖了常规结构化语言所能表达的各种算法在最坏情况下的时间复杂度

i. 赋值、比较、算术运算、逻辑运算、读写单个常量或单个变量等，只需要1个单位时间

ii. 条件语句if(C) s1 else s2中，执行时间取（C的执行时间）+（s1和s2中执行时间最长的那个时间）

iii. case语句的执行时间是所有case子句中，执行时间最长的语句块

iv. 访问一个数组的单个分量或一个结构体变量的单个域，只需要1个单位时间

v. 执行一个for循环语句需要的时间等于执行该循环体所需要的时间乘上循环的次数

vi. 执行一个当型while(c){S}循环语句或者一个直到型do{S}while(C)语句，需要的时间等于计算条件表达式C需要的时间与执行循环S体需要的时间之和乘以循

环的次数。与规则五不同，这里的循环次数是隐含的。还需要具体问题具体分析了

- vii. 对于函数调用语句，它们需要的时间包括两部分，一部分用于实现控制转移，另一部分用于执行函数本身，这时可以根据函数调用的层次，由里向外运用前六条规则层层分析直到计算出最外层的运行时间

6) 例：求连续子序列和最大正值

- i. 在线处理法，遍历+=每个输入数，<0时重置为0，>最大值时更新，时间复杂度 $O(N)$
- ii.
- iii.
- iv.
- v.
- vi.
- vii.
- viii.
- ix.
- x.
- xi. -----我是底线-----